DOCUMENT RESUME

ED 477 645 EA 032 547

AUTHOR	Brown, Richard S.; Ing, Marsha
TITLE	Exploring Sustained Improvement in Low Performing Schools. PACE Working Paper Series.
INSTITUTION	Policy Analysis for California Education, Berkeley, CA.
REPORT NO	PACE-WP-Ser-03-3
PUB DATE	2003-05-00
NOTE	16p.
AVAILABLE FROM	Policy Analysis for California Education (PACE), University of California, Berkeley School of Education, 3653 Tolman Hall, Berkeley, CA 94720-1670. Tel: 510-642-7223; Fax: 510-642-9148; e-mail: PACE123@socrates.berkeley.edu; Web site: http://pace.berkeley.edu/pace_about.html. For full text: http://pace.berkeley.edu/ wp_03-3.pdf.
PUB TYPE	Reports - Research (143)
EDRS PRICE	EDRS Price MF01/PC01 Plus Postage.
DESCRIPTORS	*Academic Achievement; Academic Standards; *Achievement Gains; *Educational Improvement; Educational Quality; Elementary Secondary Education; *Low Achievement; Performance; School Demography; School Size; Scores; *State Standards; Student Evaluation; *Student Improvement; Tests
IDENTIFIERS	*California; No Child Left Behind Act 2001

ABSTRACT

Since the approval of the No Child Left Behind Act, there have been new ways of evaluating student success. Evaluation is based on progress or improvement, not only on achievement. This report describes a study that examined student progress, or sustained improvement, in California over a 2-year period (1999-2000). In analyzing sustained improvement, the study looked at California Academic Performance Index scores in low-performing schools. The study took into account demographic and school-level variables known to relate strongly to student achievement. The report explores three issues: (1) how to conceptualize and measure sustained improvement in California public schools; (2) how the way sustained improvement is conceptualized affects schools that exhibit sustained improvement; and (3) how well predictors of student achievement relate to sustained improvement measures. The study found that the variables most commonly associated with achievement test scores were not particularly helpful in explaining sustained improvement. Social and demographic measures were not predictive of sustained improvement. Also, a potential influence on sustained improvement could be school-district support. Finally, conceptualizing and measuring sustained improvement can be approached in a variety of ways, and further research is warranted. (Contains 9 tables and 11 references.) (WFA)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 477 645

PACE WORKING PAPER SERIES 03-3

Exploring Sustained Improvement in Low Performing Schools

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY
S SMITH
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
1

Policy Analysis for California Education
University of California, Berkeley and Davis
Stanford University

Bay Area Consortium for Urban Education

May 2003

BEST COPY AVAILABLE

EA 032 547

WORKING PAPER SERIES 03-3

Exploring Sustained Improvement in Low Performing Schools

Richard S. Brown
University of California, Irvine

Marsha Ing
University of California, Los Angeles

May 2003

Policy Analysis for California Education
PACE

Table of Contents

Introduction 1

Existing Research 1

Defining Sustained Improvement 2

Describing Sustained Improvement in California Schools 1999-2001 3

Relationships Between School Demographics and Sustained Improvement 5

Summary of Findings 9

Policy Implications 10

References 10

Introduction

The recent approval of the "No Child Left Behind Act" changes the landscape for evaluating school success. In addition to requiring student testing in grades 3–8 for all students in math and reading, a major component of the new law mandates that all states determine, and schools and school districts demonstrate, "adequate yearly progress" toward state proficiency goals (No Child Left Behind Act, 2001).

> All students—regardless of race or socioeconomic status—must be held to the same academic expectations, and all students—regardless of race or socioeconomic status—must have their academic *progress* measured using a newly-refined concept of adequate yearly progress (AYP). (Keegan, Orr, & Jones, 2002)

The federal government's new approach toward evaluating public education is based on progress, or *improvement*, in student achievement, rather than on student achievement alone. But just as there is no agreement over what constitutes proficient academic performance (Olson, 2002), there is also no universal standard for measuring improvement at the school level.

In California, increases in annual Academic Performance Index (API) scores are often touted as indications of student improvement and subsequently school success. But yearly fluctuations in API scores may actually be due more to factors other than student improvement (Rogosa, 2000; Kane & Staiger, 2002). Likewise, schools that improve dramatically one year tend to improve less in subsequent years, and many actually decline.

For these reasons, it is advisable to look at student progress over a longer time frame, taking into account all the known variables that impact it. In this study, we have defined *sustained improvement* as increases in student outcomes that persist or extend beyond a single year. In addition to looking at API scores, we also consider demographic and school-level variables that are known to relate strongly to student achievement (McCullough, 2001). These include student poverty, English proficiency, school size, and teacher quality.

The following questions are explored in this report:

- How should we conceptualize and measure "sustained improvement" in California public schools?
- Does the way we conceptualize this concept impact the number of schools at each level (elementary, middle, and secondary) that exhibit sustained improvement?
- How well do predictors of student achievement relate to sustained improvement measures?

In this report, we first discuss each of these issues. Next, we present a summary of our findings. The final section considers the policy implications of this research.

Existing Research

California's Academic Performance Indicator (API) purports to measure academic performance and growth at the school level. This evolving numerical indicator ranges from 200 to 1000 and is calculated annually from various academic measures, depending upon the year. For the years analyzed here, the API was comprised primarily of the Stanford Achievement Test (SAT-9), Version 9, a nationally norm-referenced achievement test given to nearly all California school children in grades 2–11. (For more information on the API see the California Department of Education website (http://api.cde.ca.gov/.)

Gains in annual API scores for the 1999–2000 and 2000–2001 school years are documented in a series of reports prepared for the California Department of Education (CDE) (Rogosa, 2001a, 2001b). These reports suggest that persistent growth is fairly common among California elementary and middle schools, but less so among high schools. Most schools' API scores rose in both years, although the gains in 2000–2001 were typically much less pronounced than in the previous year.

For both elementary and middle schools, about half the schools showed substantial improvement in 1999–2000 (more than 10 points), and showed some improvement (at least 1 point) again in 2000–2001.

Conversely, about half the elementary and middle schools showing substantial improvement in 1999–2000 (more than 10 points) did not demonstrate API gains in the subsequent year.

Consistent improvement was much less frequent among high schools. Only about 56.5% of all high schools showed substantial improvement in 1999–2000 (more than 10 points), with 45.9% showing some improvement the following year. The rest (54.1%) either showed no improvement or lost API points in the second year.

These analyses characterize sustained improvement in a very specific manner—*substantial* improvement in the first year, followed by *any* improvement at all in the second year. Based on this definition, it may be argued that many (though not most) schools are demonstrating consistent improvement. Thus, one of the questions of this study is whether these findings are supported by research on other models of sustained improvement. This study also focuses only on initially low-performing schools—those with very low achievement rates prior to fall 1999—rather than on all public schools. This is because it is precisely the low-performing schools that most require improvement, especially sustained improvement.

Defining Sustained Improvement

What constitutes "sustained improvement" is a matter of considerable judgment and open to debate. In this report, we look at four approaches based on achievement gains in the API. Rather than looking at substantial gains in one year followed by any gains in the subsequent year, as did the California Department of Education analyses, we look at aggregated two-year gains, and achievement of specific gain amounts each year for each school. We also suggest a measure that addresses specific gain amounts for each relevant subgroup of students for each school. Lastly, we consider a more stringent improvement criterion that requires attainment of both overall school API targets and comparable improvement targets.

These measures were chosen because specified API targets—both for the school as a whole and for relevant subgroups—are the measures by which schools are typically evaluated. Simply gaining on the API may not be sufficient to demonstrate reasonable improvement, particularly for low-performing schools.

Thus, the four measures of sustained improvement suggested here include:

Outcome 1: Two-year change in API scores

This is a simple calculation of the change in API scores over two years, determined by subtracting each school's 1999 API score from its 2001 API score.

Outcome 2: Meeting school-wide growth targets for two years

Individual school growth targets vary by school and are calculated as 5% of the difference between the school's 1999 base-year API score and the statewide goal of 800, with a minimum growth target of 1 point. Schools that met or exceeded their school-wide growth targets for both years were coded as a "1". If a school failed to meet its school-wide growth targets for either or both years, they were coded as "0".

Outcome 3: Meeting comparable improvement targets for two years

Comparable improvement targets are established for each relevant student subgroup at each school. Relevant subgroups might include racial/ethnic groups, socioeconomically disadvantaged students, or English language learners. Targets are established as 80% of the school-wide growth target. For example, if the school-wide growth target is 100 API points, each relevant subgroup would need to demonstrate an 80 point API increase to meet the comparable improvement growth target criterion.

Outcome 4: Award eligibility for two years

This criterion examines a school's eligibility for Governor's Performance Awards (GPA) over two years. Based on CDE guidelines, if a school meets its school-wide and comparable improvement growth targets for *either* year, it is eligible for a GPS award. Schools that meet *both* targets for *both* years are considered to have demonstrated "sustained improvement," whereas schools that fail to meet *either* target for *either* year are not.

BEST COPY AVAILABLE

Describing Sustained Improvement in California Schools 1999-2001

The analyses presented here focus on improvement in initially low-performing public schools in California. Low-performing schools are defined as schools with a 1999 API decile rank less than or equal to 5 (Table 1). That is, these schools were those that were in the lower half of all schools in performance based on 1999 test data. The lowest decile (1) reflects the lowest 10% of all schools, the second lowest decile (2) represents the next lowest 10% of schools, and so on.

For Outcome 1, two-year change in API scores, elementary schools exhibited a much higher mean improvement (73.69) than did middle and high schools (Table 2). This is consistent with CDE reports (Rogosa, 2001a; 2001b). Overall, all three school types showed positive average changes in API scores over the two-year period for all deciles.

Over half of all elementary schools in each decile met their school-wide growth targets for the past two years (Table 3). In contrast, percentages for middle schools were considerably lower. Most significantly, few high schools met these targets. Only 4.6% of all 1999 API decile 1 high schools met school-wide growth targets for both years, and only 8.4% of high schools met school-wide growth targets for both years.

Results for Outcome 3, comparable improvement targets (Table 4), were similar to those of Outcome 2. Again, elementary schools showed sustained improvement at a higher rate than did middle and high schools. It should be noted that 51.3% of 1999 API decile 1 elementary schools met their comparable improvement targets in both years. This is in marked contrast to only 4 high schools (4.6%) in decile 1 who fulfilled the same target goal. Lower decile elementary schools met their comparable improvement targets at greater rates than higher decile schools (51.3% for decile 1 schools versus 36.7% for decile 5 schools).

By definition, schools eligible for GPA awards needed to meet their school-wide and comparable improvement targets. Once again, elementary schools performed better than middle or high schools on this fourth measure of sustained improvement (Table 5). Whereas 39% of low-performing elementary schools were awards-eligible, only 17% of middle schools and less than 3% of high schools were eligible for awards in both years.

The data indicates that sustained improvement occurs more frequently in elementary schools than in middle and high schools, which is consistent with CDE findings. While all school types exhibit positive change in their average API scores from one year to the next, significant change over consecutive years is less prevalent in middle and high schools. Moreover, as the definition of what constitutes school improvement (based on the four measures described earlier) becomes more stringent, the frequency of consistent school improvement drops significantly.

TABLE 1 Number of initially low-performing schools by type and rank (1999)

Decile	Elementary Schools	Middle Schools	High Schools
1	478	111	87
2	490	111	85
3	478	109	84
4	488	116	82
5	481	111	78
Total	**2,415**	**558**	**416**

1999 API rank ≤5
Source: California Department of Education, 1999.

BEST COPY AVAILABLE

TABLE 2 Outcome 1: Two-year change in API scores

Decile	N	Min	Max	Mean	S.D.
Elementary Schools					
1	441	-20	277	83.11	38.22
2	463	-28	202	77.08	38.52
3	461	-29	218	74.80	39.78
4	470	-49	178	69.49	36.78
5	459	-73	182	64.62	36.42
Summary	**2,294**	**-73**	**277**	**73.69**	**38.45**
Middle Schools					
1	107	-39	159	43.39	34.30
2	102	-20	146	42.88	32.40
3	105	-53	96	36.87	25.81
4	113	-51	124	35.39	28.32
5	110	-37	165	39.63	29.85
Summary	**537**	**-53**	**165**	**39.68**	**30.31**
High Schools					
1	74	-25	85	23.76	24.66
2	80	-26	77	19.90	21.08
3	78	-50	88	22.46	24.54
4	77	-40	74	20.81	25.31
5	71	-73	66	13.66	27.35
Summary	**380**	**-73**	**88**	**20.20**	**24.52**

Source: California Department of Education, January 2002.

TABLE 3 Outcome 2: School-wide growth targets

Decile	Elementary Schools	Middle Schools	High Schools
1	266 (55.6%)	26 (23.4%)	4 (4.6%)
2	268 (54.7%)	33 (29.3%)	6 (7.1%)
3	281 (58.8%)	35 (32.1%)	9 (10.7%)
4	261 (53.5%)	43 (37.1%)	11 (13.4%)
5	285 (59.3%)	53 (47.7%)	5 (6.4%)
Total	**1361 (56.3%)**	**190 (34.1%)**	**35 (8.4%)**

Source: California Department of Education, January 2002.

BEST COPY AVAILABLE

TABLE 4 Outcome 3: Comparable improvement targets

Decile	Elementary Schools	Middle Schools	High Schools
1	245 (51.3%)	25 (22.5%)	4 (4.6%)
2	236 (48.2%)	25 (22.5%)	3 (3.5%)
3	203 (42.5%)	15 (13.8%)	2 (2.4%)
4	178 (36.5%)	24 (20.7%)	3 (3.7%)
5	177 (36.7%)	24 (21.6%)	3 (3.8%)
Total	**1039 (43.0%)**	**113 (20.3%)**	**15 (3.6%)**

Source: California Department of Education, January 2002.

TABLE 5 Outcome 4: Award eligibility

Decile	Elementary Schools	Middle Schools	High Schools
1	217 (45.4%)	19 (17.1%)	3 (3.4%)
2	205 (41.8%)	21 (18.9%)	2 (2.4%)
3	191 (40.0%)	14 (12.8%)	2 (2.4%)
4	166 (34.0%)	21 (18.1%)	3 (3.7%)
5	165 (34.2%)	22 (19.8%)	2 (2.6%)
Total	**944 (39.1%)**	**97 (17.4%)**	**12 (2.9%)**

Source: California Department of Education, January 2002.

Relationships Between School Demographics and Sustained Improvement

It has been well established that school demographic characteristics, such as socioeconomic status indicators, account for much of the variation in school API scores (McCullough, 2001). Of interest here is whether the socioeconomic variables that relate strongly to API performance also predict sustained improvement.

To explore the relationship between school characteristics and sustained improvement, we selected several variables. The same variables were among those used in the California Department of Education's (2002) School Characteristic Index (SCI), and were also found by McCullough (2001) to correlate to API scores.

> "CBP's (California Budget Project) analysis found that over 80% of the variation in schools' 2000 API scores can be explained by the social and economic characteristics of a school's students, the size of the school, and the quality of its teachers." (McCullough, 2001, p. 4)

Included in our analysis were five social and economic school characteristics (Table 6). In addition, we also explored how participation in the Immediate Intervention Underperforming Schools Program (II/USP) related to sustained student improvement.

The II/USP is a state-funded program designed to provide financial assistance (over $70 million in 2000–01) to selected low-performing schools for planning and implementing school improvement plans. A cohort of 430 schools was selected for participation in this program in 1999 and an additional 430 was selected for participation the following year. It is possible that schools participating in II/USP simply have not had enough time to fully implement their school improvement plans. Consequently, it may be unreasonable to expect these schools to demonstrate sustained improvement at such an early phase.

BEST COPY AVAILABLE

A primary question of interest here is whether these school characteristics—which previously have been shown to predict initial performance (1999 API score)—also relate to sustained improvement. These relationships are expressed in terms of correlations between each of the four measures of sustained improvement (Outcomes 1–4) and these school characteristics (Tables 7-9).

For initially low-performing elementary schools, the variables that are highly related to their initial API scores are not highly related to change in API scores. There are two exceptions. Total enrollment is significantly related to three measures of sustained improvement among these schools: meeting school-wide growth targets, meeting comparable improvement targets, and award eligibility. Although it was not related to change in API scores from 1999 to 2001, there are small but significant positive relationships with the other outcome measures. These correlations suggest that for elementary schools, the larger the school, the more likely the school will show sustained improvement.

The other exception is II/USP participation. One might expect that schools eligible for II/USP participation may be less likely to meet the conceptualization of sustained positive change in API scores. Thus, the negative relationship between the predictors and the outcome measures could be expected. This is inconsistent with other research suggesting that II/USP schools have shown greater progress compared to non-II/USP schools (Farr & O'Day, 2002). It should be pointed out, however, that this study compared II/USP schools with non-II/USP schools on *annual* API gains, not on sustained improvement. This study not only uses different outcomes, but also uses all lower decile schools in the analyses, not a limited subset.

There is also a small, but significant, negative relationship between the percent of teachers with full credentials at a school and that school meeting its comparable improvement targets, award eligibility, and consistency in API scores. There is also a small, positive relationship between the percent of English language learners at a school and the school meeting its comparable improvement target and the award eligibility measure. This relationship is counterintuitive. It suggests that elementary schools with a higher percentage of English language learners met comparable improvement measure, award eligibility measure and consistency in API score change measure at a greater rate.

Also of interest is the finding that student poverty is not much related to sustained improvement, despite an exceptionally large relationship with API score (-.90). Although this predictor is related (positively) to meeting comparable improvement targets, it is not related to three measures of sustained improvement. This result contrasts with the suggestion that socioeconomically disadvantaged (SD) students are gaining at greater rates than non-disadvantaged students (non-SD). In fact, a review of median API growth for socioeconomically disadvantaged students in lower decile schools suggests that these groups progressed comparably in 2000–2001.

TABLE 6 Description of school characteristics

1. Total Enrollment	Total enrollment (1998–1999)
2. Free/Reduced Lunch	Percent of students receiving free or reduced lunch (1999–1999)
3. English Language Learners	Percent of English language learners (1998–1999)
4. Mobility	Percent of students who first attended this school in 1998–1999
5. Fully Credentialed Teachers	Percent of fully credentialed teachers (1998–1999)
6. II/USP Cohort 1	Participation in II/USP program (1999–2000)
7. II/USP Cohort 2	Participation in II/USP program (2000–2001)

Source: California Department of Education, January 2002.

BEST COPY AVAILABLE

TABLE 7 Elementary school correlations between school-level predictors and outcomes

	API	Change in API Score	School-wide Growth Targets	Comparable Improvement Targets	Award Eligibility
1. Total Enrollment		.012 (n=2291)	.044* (n=2412)	.127** (n=2412)	.102** (n=2412)
2. Free/Reduced Lunch	-.90 (n=4847)	.020 (n=2289)	-.040 (n=2410)	.050* (n=2410)	.029 (n=2410)
3. English Language Learners	-.77 (n=4847)	.085** (n=2280)	-.016 (n=2399)	.142** (n=2399)	.105** (n=2399)
4. Mobility	-.12 (n=4847)	.057** (n=2291)	.018 (n=2411)	.012 (n=2411)	.021 (n=2411)
5. Fully Credentialed Teachers	.57 (n=4847)	-.084** (n=2291)	-.031 (n=2412)	-.093** (n=2412)	-.083** (n=2412)
6. II/USP Cohort 1		.009 (n=2294)	-.056** (n=2415)	-.039 (n=2415)	-.031 (n=2415)
7. II/USP Cohort 2		-.205** (n=2294)	-.224** (n=2415)	-.265** (n=2415)	-.274** (n=2415)

*Correlation is significant at the .05 level

**Correlation is significant at the .01 level

Note: The school characteristic and API correlations are from the PSAA Technical Report 00-1 by the California Department of Education (2000). These correlations provided by the CDE are for all schools and not just for the initially low-performing schools. All figures are percents.

For decile 1 schools, median API growth in 2000–2001 for SD students was 38.1 versus 36.9 for non-SD students. For decile 2, the rates are even more similar (34.9 for SD, 34.4 for non-SD). Deciles 3 through 5 show slightly higher gains for SD students (25.3 versus 24.8; 29.7 versus 24.5; and 24.9 versus 20.7) than for non-SD students. Given the small between group differences and the substantial variation within the schools and groups, the absence of a significant correlation between student poverty and sustained improvement for this group of schools is unsurprising.

Similar to the relationships seen for elementary schools, the variables that relate to initial API scores do not relate much to change in API scores for initially low-performing middle schools. As with elementary schools, there are a few exceptions. Total enrollment is significantly related to the three of the four outcome measures. Although it was not related to award eligibility there are small but significant negative relationships with the other outcome measures. Whereas school size was positively related to sustained improvement for elementary schools, there are negative relationships between these measures at the middle school level. Thus the relationship between school size and these outcome measures appear to differ depending on the type of school.

Similar to results found at the elementary school level, there are negative relationships between II/USP Cohort 2 participation and each of the four outcome measures for middle schools. There are also negative relationships between student poverty and language fluency and meeting school-wide growth targets for two years suggesting that schools with more poverty and greater frequency of English language learners met their school-wide growth targets less frequently. This contrasts with the relationships seen at the elementary school level, where the percent of English language learners was positively related to sustained improvement on three of the four measures.

There is a positive relationship between percent of credentialed teachers and meeting school-wide and

BEST COPY AVAILABLE

TABLE 8 Middle school correlations between school-level predictors and outcomes

	API	Change in API Score	School-wide Growth Targets	Comparable Improvement Targets	Award Eligibility
1. Total Enrollment		-.112** (n=537)	-.083* (n=558)	-.111** (n=558)	-.080 (n=558)
2. Free/Reduced Lunch	-.89 (n=1121)	-.017 (n=536)	-.143** (n=557)	-.028 (n=557)	-.045 (n=557)
3. English Language Learners	-.78 (n=1121)	.023 (n=535)	-.120** (n=556)	.043 (n=556)	-.019 (n=556)
4. Mobility	-.12 (n=1121)	-.050 (n=537)	-.017 (n=558)	-.035 (n=558)	-.019 (n=558)
5. Fully Credentialed Teachers	.61 (n=1121)	-.027 (n=536)	.121** (n=557)	.041** (n=557)	.051 (n=557)
6. II/USP Cohort 1		.036 (n=537)	-.013 (n=558)	.057 (n=558)	.050 (n=558)
7. II/USP Cohort 2		-.190** (n=537)	-.125** (n=558)	-.238** (n=558)	-.217** (n=558)

*Correlation is significant at the .05 level.

**Correlation is significant at the .01 level.

Note: The school characteristic and API correlations are from the PSAA Technical Report 00-1 by the California Department of Education (2000). These correlations are for all schools and not just for the initially low-performing schools. All figures are percents.

TABLE 9 High school correlations between school-level predictors and outcomes

	API	Change in API Score	School-wide Growth Targets	Comparable Improvement Targets	Award Eligibility
2. Total Enrollment		-.074 (n=379)	-.120* (n=415)	-.112* (n=415)	-.083 (n=415)
2. Free/Reduced Lunch	-.80 (n=839)	-.054 (n=379)	-.057 (n=409)	.023 (n=409)	.015 (n=409)
3. English Language Learners	-.69 (n=839)	.006 (n=376)	-.008 (n=403)	.061 (n=403)	.078 (n=403)
4. Mobility	-.19 (n=839)	-.045 (n=380)	-.052 (n=416)	-.048 (n=416)	-.031 (n=416)
5. Fully Credentialed Teachers	.48 (n=839)	.078 (n=379)	.076 (n=414)	.058 (n=414)	.062 (n=414)
6. II/USP Cohort 1		.046 (n=380)	.050 (n=416)	.129** (n=416)	.071 (n=416)
7. II/USP Cohort 2		-.118* (n=380)	-.045 (n=416)	-.088 (n=416)	-.078 (n=416)

*Correlation is significant at the .05 level.

**Correlation is significant at the .01 level.

Note: The school characteristic and API correlations are from the PSAA Technical Report 00-1 by the California Department of Education (2000). These correlations are for all schools and not just for the initially low-performing schools. All figures are percents.

BEST COPY AVAILABLE

comparable improvement targets. It is interesting to note that this relationship is different from what is seen at other school levels.

As with elementary and middle schools, the variables that relate to initial API scores do not strongly relate to sustained improvement for initially low-performing high schools. There are few significant correlations between these school characteristics and the outcome measures. Though these characteristics are highly predictive of initial API scores, there were few significant relationships between these characteristics and the outcome measures defining sustained improvement.

School size however, is one variable that relates to the sustained improvement measures at the high school level. The relationship, however, is not the same for all school levels. For elementary schools there is a positive relationship between school size and sustained improvement. However, for middle and high schools, there are negative relationships between school size and sustained improvement. In elementary schools, larger schools were more likely to show sustained improvement. In middle and high schools, smaller schools tended to exhibit sustained improvement more than larger schools.

Summary of Findings

It appears that the variables frequently associated with achievement test scores are not particularly helpful in explaining sustained improvement (over a two-year period) or changes in test scores. We found that for all school types, the relationships between school characteristics and outcomes were extremely low. The demographic and socioeconomic variables that were highly related to *initial* API scores (in 1999) were not as strongly related to *change* in API scores (in 2000 and 2001).

It is also possible that variations in outcomes *between* schools may be better explained by variables not included in these analyses, such as school leadership, community involvement, and school culture. What is of interest here is that by recognizing the lack of effect of socioeconomic measures on sustained improvement, we can move toward identifying other potential influences not related to socioeconomic status.

Another potential influence on sustained improvement may be school district support. When we looked at Bay Area schools by district, it appeared that many of the schools showing sustained improvement were clustered in specific districts. This is an area that warrants additional research.

Conceptualizing and measuring sustained improvement can be approached in a variety of ways. In this report we present four distinct conceptualizations of sustained improvement. We then show how each impacts the number and level of schools identified as demonstrating sustained improvement. Simply looking at increases over a two-year interval in API scores will yield the largest number of schools showing sustained improvement, but it does not reflect improvement for all students. This approach also demonstrates that the magnitude of sustained improvement diminishes across the grade spans for initially low-performing schools. Overall, the largest gains were shown by elementary schools. Middle schools experienced more modest gains, and high schools substantially fewer.

While most elementary schools showed consistent increases in API scores and met their school-wide growth targets for consecutive years, most did not show consistency in meeting comparable improvement targets. The picture is less favorable for middle and high schools. Most middle schools show API improvement, but most do not consistently meet their school-wide or subgroup growth targets. High schools generally do not show consistency in improvement on any of the measures.

Social and demographic measures (student poverty, mobility, language fluency, teacher quality, and school size) were not predictive of sustained improvement.

One of the limitations of this study is that the outcomes of sustained improvement were derived from API scores. Since the composition of the API is not directly aligned to curricula in the classroom, it may not be responsive to changes at the school level. Future research should use other evaluative tools to measure sustained improvement.

This study was deliberately restricted to initially low-performing schools. As a result, these findings may not be consistent with outcomes observed in higher-performing schools. This study also covers a brief two-year time span, which may be insufficient to allow schools to enact effective changes and demonstrate improvement. More than two years of information would provide a better basis for drawing conclusions regarding consistency in school improvement. However, even two years is more informative than merely reviewing annual API changes and inferring the existence of lasting or sustainable school improvements.

Policy Implications

How "sustained improvement" is defined makes a great deal of difference in the results obtained when measuring school progress. More restrictive definitions suggest that most schools—particularly at the more advanced grade spans—simply do not demonstrate sustained improvement over the two-year period of 1999–2001.

In our study, sustained improvement was much more common in elementary schools than in middle and high schools, which is consistent with the one-year patterns observed in the CDE reports. In fact, we found a consistent progression across the grade spans; more improvement in elementary schools, less in middle schools, and still less in high schools. One policy implication of this finding is that more attention should be focused on developing programs and approaches that increase the level of improvement observed at the middle and high school levels.

Since there exists a great deal of variability in improvement that cannot be attributed to randomness—particularly at the elementary and middle school level—this points to the need for additional research. One approach would be to study those schools showing sustained improvement to identify persistent, nontransient influences on student achievement.

Another finding was that school size was positively related to sustained growth in elementary schools, while negatively related to sustained growth in the middle and high schools. This suggests a need to explore different approaches to improving student achievement in the three school types. Also, since past research has shown that reducing class size has little or no noticeable impact on student achievement (Stecher and Bohrnstedt, 2002), reducing *school* size might have a positive (although small) impact on it, at least at the middle and high school levels.

Finally, it is important to remember the distinction between achievement and improvement. As we've seen from this research, the variables associated with student achievement are not necessarily predictive of student improvement, nor do they operate the same across grade spans. If California public schools are to be judged on "adequate yearly progress" pursuant to the No Child Left Behind Act, it is wise to focus on identifying measures or practices that best relate to—and can enhance—improvement at all levels.

References

Farr, B. and O'Day, J. (2002). *Evaluation study of the Immediate Intervention/Underperforming Schools Program and the High Achievement/Improving Schools Program of the Public Schools Accountability Act of 1999.* Retrieved August 20, 2002, from http://www.air.org/pubs/Phase1ReportFinal.pdf

Kane, T.J. and Staiger, D.O. (2002). Volatility in school test scores: Implications for test-based accountability systems. In *Brookings papers on education policy 2002,* D. Ravitch, ed., pp. 235–273. Washington, D.C.: Brookings Institute Press. Retrieved April 30, 2002, from http://muse.jhu.edu/journals/brookings_papers_on_education_policy/v2002/2002.1kane.pdf.

Keegan, L., Orr, B.J., and Jones, B.J. (2002). *Adequate yearly progress: Results, not process.* Retrieved April 30, 2002, from http://www.edexcellence.net/NCLBconference/Keegan%20Orr%20and%20Jones.pdf.

McCullough, D. (2001). *California budget project budget brief: What do the 2000 API results tell us about California's schools?* Retrieved April 30, 2002, from http://www.cbp.org.

No Child Left Behind Act of 2001, Public Law 107–110, 107th Congress, 1st Session, 2002.

Olson, L. A 'proficient' score depends on geography: Achievement levels vary widely by state. (2002, February 20). *Education Week,* 21(23):1, 14, 15.

Rogosa, D.R. (2000). *Interpretive notes for the Academic Performance Index.* Retrieved on April 30, 2002, from http://www.cde.ca.gov/psaa/api/fallapi/apnotes.pdf.

Rogosa, D.R. (2001a). *Year 2000 update: Interpretive notes for the Academic Performance Index.* Retrieved on April 30, 2002, from http://www.cde.ca.gov/psaa/api/yeartwo/growth/interp2k.pdf.

Rogosa, D.R. (2001b). *Year 2001 growth update: Interpretive notes for the Academic Performance Index.* Retrieved on April 30, 2002, from http://www.cde.ca.gov/psaa/api/yeartwo/growth/interg01.pdf.

Stecher, B.M and Bohrnstedt, G.W. (Eds.). (2002). *Class size reduction in California: Findings from 1999–00 and 2000–01.* Retrieved April 30, 2002, from http://www.classize.org/techreport/year3_technicalreport.pdf.

Technical Design Group of the Advisory Committee for the Public Schools Accountability Act of 1999. (2000). Construction of California's 1999 School Characteristic Index and Similar Schools Rank. PSAA Technical Report 00-1. Retrieved April 30, 2002 from http://www.cde.ca.gov/psaa/tech0400.pdf.

Policy Analysis for California Education PACE
University of California, Berkeley and Davis
Stanford University
3653 Tolman Hall
Berkeley, CA 94720-1670

Telephone: (510) 642-7223
http://pace.berkeley.edu

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

NOTICE

Reproduction Basis

[X] This document is covered by a signed "Reproduction Release (Blanket)" form (on file within the ERIC system), encompassing all or classes of documents from its source organization and, therefore, does not require a "Specific Document" Release form.

[ ] This document is Federally-funded, or carries its own permission to reproduce, or is otherwise in the public domain and, therefore, may be reproduced by ERIC without a signed Reproduction Release form (either "Specific Document" or "Blanket").